J. Clin. Chem. Clin. Biochem.
Vol. 18, 1980, pp. 105–109

Comparison of Four Different Algorithms for the Calculation of Radioimmunoassay Standard Curves[1])

By *I. Marschner, R. Herndl* and *P. C. Scriba*

Medizinische Klinik Innenstadt der Universität München, Munich, W.-Germany

(Received July 13, 1979)

Summary: Standard curves from 150 determinations of 10 different hormone radioimmunoassays have been evaluated using the following four procedures: logit transformation as performed by *Rodbard, Healy*'s 4-component-logit, spline approximation and a linear interpolation program. The results, calculated by computer, have been compared with manual analysis as the reference method. The influence of various correction factors has been studied by artificial insertion of outliers.

Approximately two thirds of all investigated curves could be calculated with almost equal success using all 4 methods. In calculating the remaining third of the standard curves, each method showed characteristic errors which depended upon the reliability of the assay data and the kind of outliers.

The results suggest that the smoothing by spline function is the most flexible method in approximating a radioimmunoassay standard curve. In comparison with the other methods it is less influenced by random errors and fits the characteristic and symmetry of the ideal curve most exactly.

The 4-component logit which has been extended by data screening is an improvement of the original logit transformation. Certain reservations exist, however, in elimination of outliers because of the dispersion of the data points, and also because of the arbitrarily set thresholds.

Vergleich unterschiedlicher Algorithmen zur Berechnung von Radioimmunoassay-Standardkurven

Zusammenfassung: Die Standardkurven aus 150 Radioimmunoassays (RIA) von 10 Hormonbestimmungen wurden nach 4 Verfahren – nämlich der Logit-Transformation nach *Rodbard*, dem 4-Komponenten-Logit von *Healy*, der Spline-Approximation und einem linearen Interpolationsprogramm – mittels EDV berechnet und die Ergebnisse mit der graphisch manuellen Auswertung als Bezugsmethode verglichen. Durch artifizielles Einfügen von Ausreißern wurde der Einfluß verschiedener Korrekturglieder untersucht.

Etwa zwei Drittel aller untersuchten Kurven ließen sich mit den 4 Verfahren annähernd gleich gut berechnen. Das letzte Drittel führte bei den einzelnen Methoden in Abhängigkeit von der Güte der Ausgangsdaten und der Art der aufgetretenen Ausreißer zu für die jeweiligen Berechnungsverfahren charakteristischen Fehlern.

Die Ergebnisse zeigen, daß die Glättung mittels Spline Funktionen die flexibelste Methode zur Approximation einer RIA-Standardkurve ist. Sie erweist sich gegenüber fehlerbehafteten Daten stabiler als die anderen Methoden und paßt sich der Charakteristik und Symmetrie der zu berechnenden Kurven am genauesten an.

Das mit einem Datenscreening erweiterte 4-Komponenten-Logit stellt eine Verbesserung der ursprünglichen Logit-Transformation dar, wenn auch prinzipielle Bedenken gegen eine Ausreißereliminierung auf Grund der Streuung von Meßpunkten wegen der unsicheren Bestimmung der Schwellenwerte bestehen bleiben.

Introduction

Influenced by the increasing number of radioimmunological samples stemming from clinical routine and research, methods for the calculation of the assay standard curves directly from count rates have increasingly replaced the classical graphic-manual evaluation procedure. Since the first publication of *Rodbard* et al. (1), describing the logit transformation, numerous publications have been written about modifications of this approach (2, 3). Equally numerous procedures have been

[1]) Supported by the Deutsche Forschungsgemeinschaft (SFB 51) and the Bundesministerium für Forschung und Technologie.

0340-076X/80/0018-0105$2.00
© by Walter de Gruyter & Co. · Berlin · New York

described, which mostly employ geometrical functions approximating the sigmoid form of the radioimmunoassay standard curve – some with more success than others (4, 5, 6). It is notable that some of these calculating procedures show considerable deviations from the original data, especially in those cases of more than usually oscillating curves.

The aim of this publication was to compare four well known procedures for the calculation of radioimmunoassay standard curves based on different mathematical algorithms. The reference method for all data was the manual graphical approach[2]).

Materials and Methods

For the comparison three methods of broad general applicability and equal mathematical complexity were selected. All methods based on geometric functions were rejected because of their rigidity. As an additional method, a procedure was chosen which required only a minimum of mathematical calculation, in order to justify the use of the other three algorithms. In this comparison the four following methods were chosen.

Algorithms

Logit transformation (LoR)

The logit transformation (8) attempts to linearise the radioimmunoassay standard curve by a transformation of the ordinate according to the formula:

(Equ. 1) $$y = \frac{B - N}{B\emptyset - N}$$

(Equ. 2) $$\hat{y} = \text{logit } y = \log \frac{y}{1 - y}$$

(Equ. 3) working-logit: $$\hat{y} = \frac{y - \bar{y}}{y(1 - y)} + \bar{\hat{y}}$$

B = count rate of the bound fraction
B∅ = count rate of the bound fraction of the zero standard
N = non-specific binding

The variables with the bar stem from the respective previous iteration.

The mathematical part of the computer program represents a weighted regression, which gives preference to the data points with lower count rates. This corresponds to the hypothesis that precision decreases with increasing count rates (error model 2 of *Rodbard & Lewald* (8)). The straight line, representing the standard curve, is calculated from the means of the count rates, the corresponding standard concentrations and a linear equation, determining the weighting of the means. The first rough calculation of an unweighted regression line is followed by the determination of the weighting equation. In addition, the mean values of the replicates are corrected by using the so called "working logit". This iteration is done five times. The program used for our comparison was established in a modular way as published by *Rodbard & Hutt* (3). Proceeding from this logit-transformation, which alternates the rough data by means of weighting and working logit, three further modifications were developed:

I) Calculation of the standard curve, using the weighting procedure but without the working logit.

II) Calculation of the standard curve, using the working logit but not the weighting procedure.

III) Calculation of the standard curve, using neither working logit nor weighting procedure.

4-component-logit (LoH)

This procedure has been described by *Healy* (2), and uses the following equations:

(Equ. 4) $$\bar{\bar{y}} = a + b \cdot q$$ $\bar{\bar{y}}$ = count rate

(Equ. 5) $$q = \frac{z}{z + 1}$$ a is initially B∅, than optimized

(Equ. 6) $$z = e^{c - d \cdot \ln x}$$ b is initially N, than optimized

6) in 5) in 4) c, d = curve parameters, which are optimized during calculation

x = standard concentration

(Equ. 7) $$\bar{\bar{y}} = a + b \frac{e^{c-d \cdot \ln x}}{1 + e^{c-d \cdot \ln x}}$$

(Equ. 8) $$\frac{\bar{\bar{y}} - a}{b} = \frac{e^{c-d \cdot \ln x}}{1 + e^{c-d \cdot \ln x}}$$

Parameters a and b are obtained from a linear regression, setting up the parameters c and d at first arbitrarily, and then optimizing them by iteration. An outlier screening is done by statistical analysis before proceeding to the curve calculation part of the program. This eliminates replicates with too large a standard deviation as well as mean values of measure points which lie too far apart in the preliminary calculated curve. The actual criteria for the outlier screening routine are calculated for each assay individually from the standard deviations of the rough data. The program version used for the comparison was kindly supplied by the author and was able to be run on the department computer without modification.

Spline approximation

Smoothing by spline functions is a widespread method in both technology and physics for the curve fitting of data, which are affected by errors due to experimental reasons (9). In this algorithm, the weight of each measure point is determinated by the actual standard deviation of the replicates of each standard concentration (10, 11). The final curve passes closer to a mean value when the standard deviation about this point is smaller. This is influenced by a smoothing factor, which limits the sum of the distances of each mean value to the final curve, dependent upon the corresponding standard deviations (formula 11). The oscillation of the curve is minimized by reducing the square of the area under the second derivative of the curve to a minimum (formula 12). Each standard curve is composed of 3rd degree polynomial functions definable as sectors between two adjacent measure points and which are twice derivable at their connecting points. The final function is that one from all twice derivable functions g (x), which minimized condition 12 with respect to condition 11.

(Equ. 9) $$g(x) = f_i(x) \qquad x_i \leqslant x \leqslant x_{i+1} \qquad i = 0, n$$

(Equ. 10) $$f_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3$$

(Equ. 11) $$\sum_{i=0}^{n} = \left(\frac{g(x_i) - y_i}{sd_i}\right)^2 \leqslant s$$

(Equ. 12) $$\int_{x_0}^{x_n} g''(x)^2 \, dx \rightarrow \text{minimum}$$

[2]) Preliminary results have already been presented at the Annual Congress of the German Society for Endocrinology in 1975 (7).

(Equ. 13) $s = n \cdot f$; $1.5 \geq f \geq 0.2$

y_i = mean percent bound of x_i

x_i = standard concentration

sd_i = standard deviation of x_i

s = smoothing parameter

n = number of concentration steps of the standard curve

f = smoothing factor

a_i, b_i, c_i, d_i = evaluated polynomial parameters

Linear interpolation (I-Pol)

This method connects the mean values of the replicates with straight lines. It has deliberately been chosen as the simplest procedure of expressing the standard curve mathematically.

Data processing equipment

A Siemens 404/3 computer with 64 kilobytes, disc operating system, plotter, fast printer, punched tape, and punched card devices was used. All programs were written in FORTRAN IV. The rough data were taken from 125 routine assays of 10 different hormones (thyrotropin, triiodo-thyronine, thyroxin, lutropin, follitropin, somatotropin, prolactin, gastrin, insulin and argininvasopressin), and 25 standard curves from a quality control survey on the radioimmunoassay of insulin, organized by the German Diabetes Association in 1974. The number of concentration steps ranged from six to ten, that of the replicates from two to five. In most of the assays serial dilution of standards was used.

Comparison procedure

At first, each curve was drawn by hand as carefully as possible in a coordinate system with a 40 cm logarithmic abscissa (hormone concentration) and a 50 cm linear ordinate ($y = (B - N)/(B\emptyset - N) \times 100\%$). At each 3%-step on the ordinate, the corresponding x-value was read from the abscissa. These 30 data pairs were punched on papertape and stored on a disc. The 4 mathematical curve fitting procedures were carried out next using on-line plotting routines fitted into the same coordinate net. The corresponding pairs of data from these procedures and the manual reference method were listed on a fast printer, together with the percentage of differences when compared with the reference method, as well as the differences between the mathematical approaches. The maximal and mean differences between two methods were evaluated for further statistical calculations only within the steep part of the curves from 77 to 23% $B/B\emptyset$. In addition to the 150 comparisons, a further 60 comparisons were run, after artificial insertion of outliers, to test the influence of weighting routines as well as questions about the curve symmetry.

Results

Types of errors

By summarizing the differences between the results of the compared methods, the following types of errors could be characterized (fig. 1).

a) Differences in the results of two methods because of the calculation of two curves with different slopes (type 1 error, fig. 1a).

b) Correspondence of the results of two methods only in one part of the curve (type 2 error, fig. 1b).

c) Deviations due to distortion of the curve segments (type 3 error, fig. 1c) or the whole curve (type 4 error, fig. 1d), resulting from outliers with small or large standard deviations, respectively.

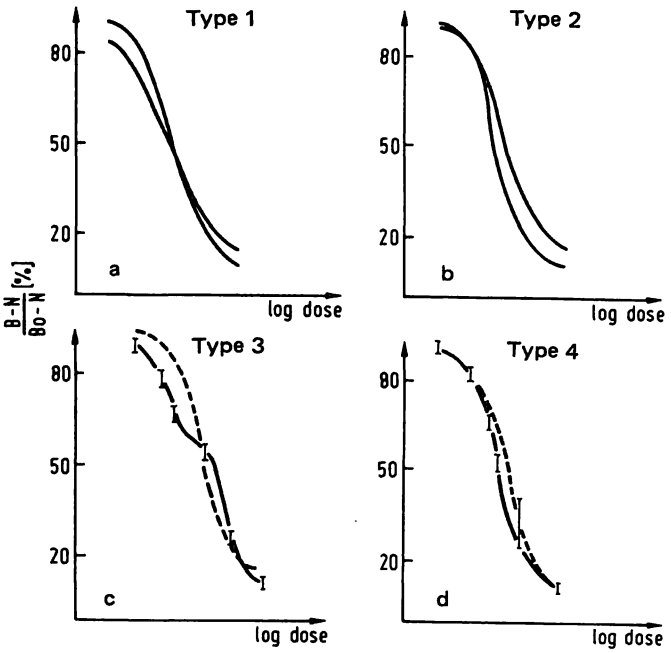

Fig. 1. Schematic representation of the 4 occurring types of error.

In all cases with more than 15% maximum error – shown in 25% of all comparisons – it was possible to use this scheme. Table 1 shows a subdivision of those cases in which deviations of more than 15 or 20% were found. The differences between the 4 mathematical methods and the reference method, as well as those between the mathematical methods, including the types of errors, were listed.

Results of the comparison between the mathematical methods and the reference method

Table 1 shows that the spline-approximation has the smallest differences when compared with the manual

Tab. 1. Subdivision of those cases, in which errors greater than 15% were found.

N = 150 Comparison of	Percentage of cases with maximum deviation		Types of errors			
	greater than 20%	greater than 15%	1	2	3	4
Reference method against spline	4	12	2	2	–	2
Reference method against LoR	12	30	10	1	–	4
Reference method against LoH	6	16	4	4	–	–
Reference method against I-Pol	24	40	–	–	–	–
Spline against LoR	8	24	14	4	3	3
Spline against LoH	6	14	6	5	2	1
LoR against LoH	5	18	4	13	–	1

approach (12). Differences larger than 15% hormone concentration were found in 12% of all cases. With the *Healy* method (LoH) however, these errors occurred in 16% and with the logit program of *Rodbard* (LoR) in 30% of the comparisons. The division into types of error showed that in most of the cases LoR had type 1 error, i. e. the calculated curve is in 9 of 10 cases less steep, and in the remaining case steeper, than the reference curves.

In the part of the standard curve examined, the difference between hand and spline was in 85% of the cases smaller than between hand and LoR. In only 6% of the cases was the reverse seen. In 86% of the cases the difference between hand and LoH was as large as the difference between hand and LoR. Only in the remaining 14% was LoH an improvement.

The linear interpolating program (I-Pol) showed in most of the cases an oscillating course around the other curves, and the results were scattered within the range of the other methods. In 60% of the comparisons, the differences between I-Pol and the reference method, or the other computerized methods represented less than 15% of the hormone concentration.

Comparison between the spline-approximation and the logit-modifications

By comparing only the mathematical methods, the greatest differences occurred between spline and LoR. In 24% of the cases, differences of more than 15% hormone concentration were found. A mean difference within the steep part of the curves of 6% hormone concentration was observed between spline and LoR. The differences between the original data and the ordinate values of the calculated curves were in 95% of the comparisons smaller with spline than with LoR, and in 5% of the comparisons they were equal (13). In most of the cases type 1 errors were found with LoR giving the flatter curve. By comparing LoR with LoH we found deviations greater than 15% in all cases where the *Healy* program eliminated standard measure points with its outlier screening routine. Here in most of the cases type 2 errors were found. In 76% of the comparisons, coincidence of LoR and LoH was found to be within 2% hormone concentration.

For all 4 investigated algorithms it was found that the results, calculated by one of the methods, were always within a one standard deviation range of the other three methods.

In contrast to *Rodbard & Hutt* (3), no significant correlation between the standard deviations of the measure points and their count rates was demonstrated.

Influence of the correction-routine of the LoR-method

In this part of the comparison, an attempt to quantify the influence of the working logit and the weighting routine of the LoR-program was made using the data of the foregoing comparisons, and artificially changing some data to test the influence of different kinds of outliers. The differences between the four modifications were up to 28% hormone concentration, and increased with decreasing precision and accuracy of the count rates.

The deviations caused by not using the weighting routine were shown to be less marked than the deviations caused by not using the working logit. In 65% of the comparisons, differences of less than 6% hormone concentration were found. The modifications using the working logit (LoR and II)) showed the smallest differences when compared with both the curves calculated by spline functions as well as the reference method.

Discussion

The results show good agreement in two thirds of the compared curves (within 10% hormone concentration) between the spline, LoR, and LoH algorithms and the graphical approach. The remainder however show rather surprising differences. Some of these differences are caused by replicates with large standard deviations. The artificial introduction of such an outlier into a standard curve, which originally gave almost identical results with all methods using the original data, led to different influences on the curves calculated by LoR, LoH, and spline. The spline approximation was not affected as the large standard deviation of that measure point reduced its weight considerably. If the LoH method rejected this standard point, the influence on the curve depended on whether the point was at the end or in the middle of the standard curve. If it was in the middle part, the curve was not affected. However, if it was in the higher or lower dose range, a type 2 error resulted. If the *Healy* method did not reject this standard point, the curve was altered in the same way as with the LoR program. If the outlier was in the middle range of the curve, a parallel deviation was found corresponding to error type 4. If it was in the upper or lower part of the curve, deviations or error type 2 were seen.

Outliers with small standard deviation (e. g. a wrong standard dilution) had the same influence on the LoH and LoR programs as those with large standard deviation. In this case, however, the weighting routine of the spline functions (formula 11) has negative consequences (error type 3). The standard curve oscillated at this point, but was not affected in the other curve segments.

A further reason for deviations between the spline, LoH and LoR is that the logit modifications always calculate curves with symmetric character with the inflection point at 50% B/BØ, which does not necessarily correspond with the radioimmunological system. The inflection points of the curves, calculated by spline functions were between 45 and 56%, thus showing rough agree-

ment with the theoretical inflection point. Only in a few assays, however, do the count rates make a symmetrical curve possible.

The analogous deviation of LoH and LoR in the cases of outliers which are not eliminated by LoH, together with the similar standard curves calculated by these algorithms, are due to the algebraic equivalence of the formulas, on which both methods are based.

Logit curve:

(Equ. 14) $\hat{y} = u + v \cdot \ln x$

Logit transformation:

(Equ. 15) $\hat{y} = \ln \frac{\overline{y}}{1 - \overline{y}} = \text{logit}\, \overline{y}$

(Equ. 16) $\ln \frac{\overline{y}}{1 - \overline{y}} = u + v \cdot \ln x$

(Equ. 17) $\frac{\overline{y}}{1 - \overline{y}} = e^{u+v \cdot \ln x \cdot}$

(Equ. 18) $\overline{y} = (1 - \overline{y}) \cdot e^{u+v \cdot \ln x}$

(Equ. 19) $\overline{y}(1 + e^{u+v \cdot \ln x}) = e^{u+v \cdot \ln x}$

(Equ. 20) $\overline{y} = \frac{e^{u+v \cdot \ln x}}{1 + e^{u+v \cdot \ln x}}$

The conversion of the logit transformation (formula 16) shows in formula 20 the identity with formula 8.

LoH and LoR do not give identical curves, because the parameters a and b, which are originally the values of N and (BØ - N), are recalculated and corrected according to the other data points. Only in a few cases after the calculation procedure do these values remain identical with the original data. The LoR program also changes the original data through using the working logit procedure and the weighting routine. The similarity of the calculated curves in 71 of 150 comparisons, despite the mathematical differences, led to the program modifications as described above.

The data show that a difference of 15–28% hormone concentration occurred only in the case of insufficient accuracy and precision of the original data (count rates). In these cases, the influence of the working logit subroutine was seen as a useful correction. The influence of the weighting procedure was found to be ambiguous, as it led to an improvement in some cases of outliers, but to errors in others. Normally such a correction lies within the expected error range. It was not necessary to differentiate procedures of weighting routines.

None of the three programs LoR, LoH, and spline approximation, which are equivalent in calculation time and storage capacity, guarantees reliable elimination of those outliers with small standard deviation. It is therefore urged that a graphic on-line representation of the standard curve on the plotter, display, or the fast printer should be made, and manual correction of outliers carried out if necessary.

References

1. Rodbard, D., Bridson, W. & Rayford, P. L. (1969), J. Lab. Clin. Med. (St. Louis) *74*, 770–781.
2. Healy, J. J. R. (1972), Biochem. J. *130*, 207–210.
3. Rodbard, D. & Hutt, D. M. (1974), in: Radioimmunoassay and related procedures in medicine, vol. 1, p. 165–192. IAEA, Vienna.
4. Täljedal, I. B. & Wold, S. (1970), Biochem. J. *119*, 139–143.
5. Vivian, S. R. & LaBella, F. S. (1971), J. Clin. Endocrinol. *33*, 225–233.
6. Vogt, W., Popp, B. & Knedel, M. (1973), this J. *11*, 438–445.
7. Herndl, R. & Marschner, I. (1975), Acta Endocrinol. (Kbh.) Suppl. *193*, 117.
8. Rodbard, D. & Lewald, J. E. (1970), Acta Endocrinol. (Kbh.) Suppl. *147*, 79–103.
9. Reinsch, C. H. (1967), Numer. Math. *10*, 177–183.
10. Marschner, I., Erhardt, F. & Scriba, P. C. (1974), in: Radioimmunoassay and related procedures in medicine, vol. I, p. 111–122. IAEA, Vienna.
11. Marschner, I., Dobry, H., Erhardt, F., Landersdorfer, T., Popp, B., Ringel, C. & Scriba, P. C. (1974), Ärtzl. Lab. *20*, 184–191.
12. Sinterhauf, K., Müller, T., Spira, H. J., Müller, D. & Lommer, D. (1975), Acta Endocrinol. (Kbh.), Suppl. *193*, 115.
13. Schoeneshoefer, M. (1975), Acta Endocrinol. (Kbh.), Suppl. *193*, 116.

Dr. med. Ingo Marschner
cand. med. Rudolf Herndl
Prof. Dr. med. P. C. Scriba
Medizinische Klinik Innenstadt
der Universität München
Ziemssenstraße 1
D-8000 München 2